Four Years of Pandemic-Era Emergency Licenses: Retention and Effectiveness of Emergency-Licensed Massachusetts Teachers Over Time

Ben Backes
James Cowan
Dan Goldhaber
Roddy Theobald

April 2024

WORKING PAPER No. 299-0424

CALDER
National Center for Analysis of Longitudinal Data in Education Research

AIR®

Four Years of Pandemic-Era Emergency Licenses: Retention and Effectiveness of Emergency-Licensed Massachusetts Teachers Over Time

Ben Backes
American Institutes for Research / CALDER

James Cowan
American Institutes for Research / CALDER

Dan Goldhaber
American Institutes for Research / CALDER
University of Washington

Roddy Theobald
American Institutes for Research / CALDER

Contents

Acknowledgments

This work was supported by a contract from the Massachusetts Department of Elementary and Secondary Education (DESE) to the American Institutes for Research (AIR). DESE reviewed a draft of this paper to ensure that confidentiality agreements regarding personally identifiable information were upheld, but AIR retained final editorial control. We thank Claire Abbott, Elizabeth Losee, Elana McDermott, and Aubree Webb for helpful discussion and comments.

CALDER working papers have not undergone final formal review and should be cited as working papers. They are intended to encourage discussion and suggestions for revision before final publication. Any opinions, findings, and conclusions expressed in these papers are those of the authors and do not necessarily reflect the views of our funders or the institutions to which the authors are affiliated. All errors and opinions are our own.

CALDER • American Institutes for Research
1400 Crystal Drive 10th Floor, Arlington, VA 22202
202-403-5796 • www.caldercenter.org

Four Years of Pandemic-Era Emergency Licenses: Retention and Effectiveness of Emergency-Licensed Massachusetts Teachers Over Time

Ben Backes, James Cowan, Dan Goldhaber, & Roddy Theobald

CALDER Working Paper No. 299-0424
April 2024

Abstract

Most states responded to the onset of the pandemic by temporarily granting teachers Emergency licenses. These licenses allowed teachers to work in classrooms without passing the typical licensure exams. Since then, several states have extended their use of Emergency licenses, raising questions about how these policies impact the composition of the teacher workforce and student outcomes. In this paper, we examine the result of these policies using data on multiple cohorts of Emergency licensed teachers (ELTs) who taught in Massachusetts between 2021 and 2023. We find that ELTs were slightly more likely to remain in the same school and in the teaching workforce than teachers from other entry routes. However, ELTs' students scored significantly lower on standardized tests in math and science than other students in the same school and same year. Our findings are at odds with earlier, more positive assessments of Emergency licensure in Massachusetts. Our updated results appear to be driven by more recent cohorts of ELTs, rather than the teachers who received Emergency licenses at the start of the pandemic. Overall, this study suggests policymakers should be cautious when drawing sweeping conclusions about the impacts of teacher licensure based solely on the earliest cohort of teachers who obtained pandemic-era Emergency licenses.

1. Introduction

As testing centers and schools closed in the spring of 2020, most states temporarily waived their requirements for licensing new teachers (DeArmond et al., 2023). Instead, these teachers received temporary licenses that allowed them to work in schools without passing licensure tests or completing other requirements, like student teaching; these licenses were temporary because states typically assigned them an expiration date. Nationwide, some estimates suggest that these types of policies could have affected around 100,000 graduates of traditional teacher preparation programs (and more nonprogram completers) who were eligible to start teaching in fall 2020 (DeArmond et al., 2023). Early research on teachers who entered the profession with Emergency licenses, or Emergency-licensed teachers (ELTs), in Massachusetts and New Jersey during the pandemic found many of these teachers were no less effective than traditionally licensed teachers (Backes & Goldhaber, 2023; Bacher-Hicks et al., 2023).

The encouraging results about pandemic-era Emergency licenses have, in some quarters, reignited long-standing debates about the value of licensure standards. For decades, states have set standards for content knowledge through their licensure testing systems, providing a minimum bar for entry into the profession. For just as long, critics have argued that licensure tests do little to guarantee quality and may prevent individuals who would become effective teachers from entering the profession. More recent critics have also noted that licensure tests also may, because of differential pass rates by race and ethnicity, contribute to the underrepresentation of Black and Hispanic teachers in the workforce and widen the "diversity gap" between students and their teachers (National Academies of Sciences, Engineering, and Medicine, 2020).

The early evidence on ELTs during the pandemic as well as these prior criticisms of tests have revived arguments that licensure requirements could be reduced or eliminated without harming teacher productivity (Aldeman, 2024a, 2024b; Yglesias, 2024). Aldeman (2024b), for instance, calls on state policymakers “to eliminate barriers to *entry*, put more of the screening responsibility on schools, where it rightly belongs, and only grant full licenses to teachers who have demonstrated a track record of performance in the classroom”. All this comes at a time when many states are actively reassessing how they determine eligibility for teaching positions (Backes & Goldhaber, 2023; Lambert, 2020; Partelow et al., 2017; Skinner et al., 2020).

In this paper, we inform these debates by examining the composition, retention, and impact of ELTs in Massachusetts over the course of the pandemic. Consistent with prior work inside and outside of Massachusetts, we find that ELTs are disproportionately more likely to be teachers of color and are more likely to be assigned to classrooms with higher shares of students of color and/or higher levels of economic disadvantage. Also consistent with prior work in Massachusetts, we find that ELTs are slightly more likely to remain in the same school and in the teaching workforce relative to other teachers with similar levels of experience and in similar teaching contexts. In other words, when states reduce barriers to entry, it appears to make teaching more accessible to candidates who are more racially and ethnically diverse.

When we look at the impact of all cohorts of ELTs, however, we find that ELTs are significantly less effective at improving student outcomes on math and science tests—by about 0.02 standard deviations—than similar teachers who entered through other routes. ELTs also receive lower teacher evaluation ratings than other teachers. On closer examination, these negative impacts are driven primarily by teachers in earlier grades in math (Grades 4–6) and by ELTs who received their first Emergency license after the initial wave of Emergency license recipients (i.e., after December 2020). We do not find any statistically significant impacts of ELTs on English language arts (ELA) test scores, nontest outcomes, or student perceptions of school climate. When we look more closely at different cohorts of ELTs, we also find that later cohorts of ELTs had less formal contact with the teacher licensure pipeline. Among the earliest Emergency license applicants, over 60% had already passed their first set of licensure tests, and a quarter had previously enrolled in a teacher preparation program. In the latest Emergency license cohort, however, the rates of prior engagement in the teacher pipeline had declined by about 40% among newly minted ELTs. The key implication is that policymakers should be cautious about drawing sweeping conclusions from the positive results associated with the earliest cohort of pandemic-era ELTs. Additional years of data suggest that selection into this licensure route has changed significantly since the beginning of the pandemic.

2. Background and Massachusetts Policy Context

When testing centers closed at the onset of the pandemic, prospective teachers were unable to take the exams they needed to get licensed. Nationwide, all but two states responded by waiving at least some of their licensure requirements (DeArmond et al., 2023). Many created temporary licenses or broadened existing emergency licenses. These adaptations allowed new teachers to enter classrooms and defer fulfilling the standard licensure requirements until a future date.

Massachusetts authorized the creation of its Emergency license shortly after the start of the pandemic. It issued its first Emergency licenses in June 2020, and it stopped issuing them on November 7, 2023.[1] Massachusetts's Emergency licenses were initially valid for one calendar year, but teachers could extend them each year by demonstrating that they were making progress toward obtaining a regular license. If they had taken one of the state's Massachusetts Tests for Educator Licensure (MTEL) or enrolled in an approved teacher preparation program, for

[1] https://www.doe.mass.edu/news/news.aspx?id=27102

example, they could continue working on an Emergency license.[2] For teachers who received Emergency licenses between June 2020 and December 2021, their credentials are scheduled to expire at the end of the 2023–24 school year. However, the state has given them the option to extend their Emergency license until the end of the 2024–25 school year.[3] To remain teaching after that final deadline, teachers will need to obtain a non–Emergency license; for most teachers, that will require passing an MTEL subject test.

2.1 Prior Literature

The vast majority of states require prospective teachers to first pass basic skills and subject-matter tests prior to becoming licensed to teach.[4] These requirements are intended to ensure that teachers meet a basic level of competency and subject-matter knowledge before they are allowed in the classroom; prior work across multiple states and licensure tests have found that these tests are predictive of eventual teacher performance, especially in math (e.g., Cowan et al., 2023, Clotfelter et al., 2007, Goldhaber & Hansen, 2010).[5] However, other research has also raised the question of whether the existence of high-stakes licensure testing actually improves the pool of entering teachers (Angrist & Guryan, 2004; Goodman et al., 2008; Podgursky, 2005). Because licensure testing acts as a screen on potential teachers and imposes additional costs, some effective teachers may be prevented from entering the profession, making the theoretical effect of licensure requirements unclear (Angrist & Guryan, 2008). These concerns are heightened because of differences in pass rates by race and ethnicity among teacher candidates

[2] The full list of possible actions includes taking MTEL (either the Communication and Literacy Skills Test or a subject test), enrolling in an approved preparation program, or completing coursework within an approved preparation program. Individuals applying for extension also must demonstrate participation in or completion of an induction program.
[3] This final extension year was announced in February 2024: https://www.doe.mass.edu/news/news.aspx?id=27319.
[4] NCES SER data table 3.1: https://nces.ed.gov/programs/statereform/tab3_1.asp.
[5] For a review, see Goldhaber, 2011.

(Goldhaber & Hansen, 2010; Nettles et al., 2011). In light of a growing awareness of disparities between the demographic composition of teachers and students, many states are reconsidering the role that licensure policies play in maintaining a high standard of teaching across all candidates (Lambert, 2020; Partelow et al., 2017; Skinner et al., 2020).

To our knowledge, only two studies measure the impact of temporary pandemic-era licensed teachers on student outcomes (Bacher-Hicks et al., 2023; Backes & Goldhaber, 2023). One of these studies focuses on early Emergency license recipients in Massachusetts (Bacher-Hicks et al., 2023). Consistent with the analysis in this paper, Bacher-Hicks et al.'s (2023) work finds that early Emergency license recipients in Massachusetts were no less effective, at least as measured by teacher contributions to student test outcomes (i.e., student growth percentiles), and more diverse than the existing teaching workforce. The second study (Backes & Goldhaber, 2023) examines a similar pandemic-era policy in New Jersey. Again, consistent with the findings in this paper, the New Jersey study finds increases in the ethno-racial diversity of the novice teacher workforce. In addition, it finds no statistically significant differences in student test scores or teacher evaluation ratings.

Both studies help us understand the early impact of Emergency license policies, but they are limited in important respects. For example, Bacher-Hicks et al. (2023) only use data from 2021–22 on ELTs in their first year of teaching. Since then, successive waves of teachers have entered Massachusetts classrooms on Emergency licenses. The study's student outcomes are also limited to test outcomes in math and ELA. The New Jersey study's sample is also limited to a cohort of just 229 unique teachers who obtained an Emergency license in the state (and only 58 in tested grades and subjects). In this study, we build on these prior studies by considering a larger sample of ELTs – and thus improving the precision of our estimates – and investigating a broader range of student outcomes, including test outcomes in math, ELA, and science, as well as nontest outcomes and student climate views.

2.2 Theoretical Model

We present a simple model of teacher licensure decisions to frame our discussion of how selection into the Emergency license may have changed over the course of the pandemic. In the spirit of Spence (1973), we sketch out a signaling model to motivate the selection issues in the Emergency license literature and our investigation. Because the main innovation of an Emergency license is the ability to obtain a teaching license without passing licensure tests – i.e., while Emergency license can be obtained without formal teacher preparation, Massachusetts already offers another license for candidates who have not completed an educator preparation program – we focus on a prospective teacher's decision to complete licensure requirements. We begin with the assumption that teacher candidates face costs c for completing licensure requirements that depend on their teaching skills h:

$$c = {}^{a}/_{h}$$

for some value of $a > 0$. The costs may include time spent studying, testing fees, and retakes. We assume that any candidate could pass the tests, although the costs of doing so may be prohibitive. In our model, the benefit of passing the licensure test is that it allows a candidate to earn a non-Emergency license and signal their teaching skills to potential employers.

Employers observe the licensure outcome and form an expectation about a teacher's effectiveness when making offers. We take employers' information as given and focus on the decision of individual teachers on whether to pursue licensure. Teachers have preferences over receiving a job offer. We normalize teacher utility to be u_1 if they work in teaching and 0 otherwise. Teachers choose to complete the testing requirements if the expected utility of doing so exceeds the testing costs. We assume that, in the absence of the Emergency license, the probability of employment for candidates who pass the test is p_1 and 0 otherwise. When the Emergency license is available, teachers can become employed without passing the test. We assume there is still some signaling value for earning the traditional license, so the probability of employment depends on license type. Traditionally licensed teachers are employed with probability p_1^e, and Emergency licensed teachers are employed with probability p_0^e, with $p_0^e < p_1^e$.

The teacher's decision is illustrated in Figure 1. Initially, teachers complete the licensure requirements when $p_1 u_1 > {}^{a}/_{h}$. This is satisfied for teachers with $h > h^*$. With Emergency licenses, teachers complete the requirements when $(p_1^e - p_0^e)u_1 > {}^{a}/_{h}$, which is satisfied for teachers with $h > h^{**}$. This generates three groups of teachers. The first group, in region A, are those with $h \leq h^*$, who enter only under the Emergency license policy. The second group, in region B, have $h^* < h \leq h^{**}$ and switch from regular to Emergency licensure when it is offered. The final group, in region C, have $h > h^{**}$ and complete the licensure requirements regardless.

There are two key takeaways from this model. First, the sample of ELTs includes both entrants, whose employment prospects are positively affected by the presence of the Emergency license, and inframarginal candidates, who only select into the Emergency license because the incremental signaling benefits of passing the tests are less than their costs. These latter candidates would have chosen to pursue standard licensure in the absence of the Emergency licensure policy. As a result, examining differences in teacher effectiveness across Emergency licensure groups is not *necessarily* indicative of the effectiveness of candidates who would enter the profession only in the absence of the traditional licensure requirements. Because our estimates of the effectiveness of ELTs include some teachers who would have found employment regardless, they do not identify the effectiveness of the new entrant group.

Second, the equilibrium is sensitive to changes in beliefs about the probability of employment from completing different licensure requirements. When there is little signaling benefit to obtaining the full license, more teachers will opt into the Emergency license. This arguably may have been the case early in the pandemic, when testing centers were closed and all parties knew that candidates had little option but to pursue an Emergency license as a temporary measure. However, if beliefs about Emergency license holders change over time (e.g., if employers see them as lower quality candidates), then the decisions of individual candidates to pursue Emergency licenses will change as well. An implication is that declining effectiveness among ELTs over the course of the pandemic could be driven entirely by changes in the composition of those candidates who would have previously obtained licensure (those in region B) rather than any change in the effectiveness of the entrants (those in region A).[6]

Finally, we note that one limitation of this partial equilibrium sketch of teacher licensure decisions is that it does not capture potential effects operating through changes in teacher supply. Increasing the supply of teachers may be beneficial even if new entrants are less effective on average if schools would have resorted to more costly measures (e.g., larger class sizes, long-term substitutes) in their absence.

---

[6] There may also be compositional changes in the group of potential entrants (region A) if more individuals became aware of the Emergency license and thus considered teaching as time passed, which may also explain differences in effectiveness across cohorts.

3. Data

Our study sample is anchored by teachers who found employment in Massachusetts in the 2020–21 through 2022–23 school years.[7] Teacher employment and licensure data are obtained through matched state records. Because we are interested in how modifications to licensure requirements affected initial routes into teaching, we identified ELTs based on whether their first teaching license was an Emergency license.[8] Using the data of first-license receipt, we also construct indicators of whether an individual held a paraprofessional role or was employed in any capacity in Massachusetts public schools prior to earning their license, whether an individual enrolled in a teacher preparation program prior to earning their license, and whether an individual took licensure tests prior to earning their Emergency license. For the latter, we construct separate indicators for whether the individual took the MTEL Communication and Literacy Skills Test (CLST) and whether the individual took an MTEL subject test prior to earning their Emergency license. In addition, using teaching assignment codes, we construct separate indicators for whether an individual taught a core subject course, a core subject course for students with disabilities, and a core subject course for English learners.

We calculate a teacher's years of experience as the difference between the current school year and their first-observed school year teaching.[9] In the regressions described below, we

[7] We explore alternate specifications that began with 2016–17 because that was the first year of the current statewide testing regime in Massachusetts after the state moved away from PARCC. We sometimes include these additional years of data to compare, for example, novice teacher composition before and after the pandemic (e.g., Table 1). Results are similar when using post-pandemic years only; see Table 4.
[8] This is the vast majority of Emergency licenses granted (Appendix Table B1). Individuals who earned an Emergency license after already earning another teaching license received an Emergency license in a Specialist field, such as an elementary school teacher already licensed in Elementary prior to the pandemic who later added on an Emergency license in Reading.
[9] Available employment records go back to the 2007–08 school year, meaning that some individuals initially observed as teachers in 2007–08 were actually teaching before. However, because we binned experience and the highest bin is more than 10 years of teaching experience, observing their true years of experience would not affect how these older cohorts of teachers are classified in regression models.

control for binned years of experience as follows: first year of teaching, second year of teaching, third or fourth year, fifth or sixth year, seventh to 10th year, and more than 10 years of experience. We also construct teacher retention measures using employment records that describe whether an individual was found in the same school in the following year or in the teaching workforce in the following year. In order to calculate retention following the 2022-23 school year for inclusion in the analyses, we use an October 2023 staffing snapshot (i.e., from Fall of the 2023-24 school year).

To estimate teacher performance, we match teachers to the outcomes of the students they teach. Because recent research has emphasized the multidimensional nature of teacher effectiveness (Jackson, 2018), including research in Massachusetts (Backes et al., 2022, 2023), we used three types of student outcomes. First, we use standardized test scores for mathematics and ELA courses in Grades 4–8 and 10 along with end-of-grade tests in Grade 5 and Grade 8 in science and end-of-course tests in biology and introductory physics (taken in Grades 9 or 10), all standardized within test and school year. For these outcomes, we restrict the sample to students in tested grades and subjects (i.e., ELA, math, and science in the relevant grade/subject), and in some models we estimate subject-specific test impacts.

Second, we construct additional measures of teacher effectiveness using nontest student outcomes used in prior research (Backes et al., 2023; Backes & Hansen, 2018; Gershenson, 2016; Jackson, 2018), including absences, disciplinary infractions, grade progression, and grade point average. As in prior work, we construct a single nontest factor using factor analyses of these four outcomes (excluding grade point average in elementary school where it is not available). We should note that prior research in Massachusetts has found that these nontest student outcomes capture teacher effects on long-run student outcomes, such as high school completion and college enrollment, that are distinct from teacher effects operating through test scores (Backes et al., 2023). For nontest outcomes, we estimate teacher impacts in Grades 4–11 in the four core academic subjects: ELA, math, science, and social studies.

Finally, we use the Views of Climate and Learning survey from Grades 4, 5, 8, and 10 conducted in the 2018, 2019, 2021, 2022, and 2023 school years to capture school climate.[10] Recent work in Massachusetts has found that teachers systematically influence students' views of school climate and that students of teachers with high climate value added tend to have better test and nontest outcomes, especially in elementary school (Backes et al., 2022). As with the other nontest outcomes, the climate surveys are student-level measures. We include the four academic subjects and surveyed grades in our analysis sample for school climate.

Due to the pandemic, we do not use student test data from the 2019–20 school year or student nontest data from the 2020–21 school year as outcome measures (the latter due to unreliable measures of absences and suspensions caused by virtual schooling). To avoid dropping excessive amounts of data in these years, we impute lagged test scores and lagged nontest outcomes for right-hand-side variables using the relationship estimated by a regression of test or nontest outcome on cubic functions of prior test and nontest outcomes. For example, for a fifth grader in the 2020–21 school year, prior test scores from 2019–20 are imputed based on the student's test and nontest results from third grade in 2018–19. Again, we perform imputation only for right-hand-side variables and do not impute outcomes.

In addition to student outcomes, we also use annual performance evaluations as another measure of teacher performance. Prior work has found that performance evaluation ratings predict student achievement on both mathematics and ELA tests (Cowan et al., 2020). As with

[10] With the exception of Grade 4 in 2018, which was not offered in this first year of administration.

other outcomes, we standardized these measures to be mean 0, standard deviation 1 in each school year.

4. Composition of Emergency-Licensed Teachers Over Time

The Appendix Table B1 displays raw counts of Emergency licenses. Nearly 20,000 unique individuals received Emergency licenses between the onset of the pandemic and the end of the policy in November 2023. Nearly 16,000 of these received teaching (as opposed to, for example, administrative) licenses, and about 15,000 of these individuals did not already hold a non−Emergency teaching license. As of the 2022−23 school year, about half of teachers whose first teaching license was an Emergency license have been hired into teaching positions.

Figure 2 displays a histogram of Emergency licenses granted by month and year. The single month with the most Emergency licenses granted was June 2020, the first month that the licenses were offered. In general, license receipt peaks in the summer months between school years. Based on these patterns and Massachusetts policy, we divide Emergency license recipients into three cohorts. The first cohort received their first license in the 2020 calendar year (June through December). As shown in Figure 2, this was the first wave of license recipients. The second cohort received their Emergency license during the 2021 calendar year (note that these two cohorts are eligible for extensions through 2025 and obtained their licenses before a pause in the policy after December 2021). The last cohort includes teachers who received their first Emergency license in 2022 or later.[11] Finally, we display the number of individuals earning their first license by license type and calendar year in Appendix Figure B1. The rise in Emergency licenses after the onset of the pandemic was associated with decreases in newly-licensed

[11] Because the bulk of licenses are granted over the summer between school years, results are not sensitive to the choice of month used to divide cohorts (i.e., instead of between December and January). In Figure A1, we plot month-specific estimates by interacting month with ELT.

individuals with both Initial (for those who had completed an educator preparation program) and Provisional (for those who did not) licenses, while the total number of newly-licensed individuals increased after the onset of the pandemic. This suggests that in addition to diverting some prospective teachers away from Initial and Provisional licenses, many individuals who earned an Emergency license would not have chosen to become licensed to teach in the absence of the Emergency license option.

Table 1 explores differences across the Emergency license cohorts. Cohort 1—the group who received an Emergency license at the onset of the pandemic between June and December of 2020—is observationally distinct from the others. Relative to the other cohorts, ELTs from Cohort 1 were more likely to have an assignment teaching core subjects to students with disabilities and were much more likely to have already taken the MTEL prior to receiving their Emergency license. They were also much more likely to have passed the CLST portion of the MTEL prior to earning their license than later cohorts and to have enrolled in a teacher preparation program prior to earning their license. However, only 27% of this first cohort had passed an MTEL subject test prior to earning their Emergency license, despite being much more likely to have attempted a subject test. Taken together, the initial Emergency license cohort was likely to have been interested in teaching (they had high rates of being a paraprofessional and taking MTEL CLST) but was not yet licensed because they had not passed their MTEL subject test(s).

The cohorts appear to have varied in other ways as well. For example, the average standardized evaluation rating for non-ELTs who earned their license in 2017 or later is −0.34 compared with −0.56 for the first cohort, −0.66 for the second, and −0.77 for the third cohort (these numbers are all negative because newer teachers tend to have lower evaluation ratings).

The final rows of the table show counts of unique ELTs and total Emergency licenses granted. The earlier Emergency license cohorts have had more time to search for employment, meaning that a higher percentage of them are observed in the teacher data. For example, 2,586 of 3,761 Cohort 1 Emergency license recipients obtained a teaching job by the 2023 school year (69%) compared with 3,960 of 6,459 for the most recent cohort (61%). It is thus possible that patterns for more recent cohorts will change as data from additional school years become available and more prospective teachers from these cohorts find teaching positions.[12] Consistent with prior work in Massachusetts and New Jersey, we also found that all ELTs are more likely to be teachers of color, with Black teachers ranging from 11%–14% of ELTs across cohorts and Hispanic teachers constituting 12% compared with 3% (Black) and 4% (Hispanic) of non-ELTs who recently received their teaching license.

Table 2 displays descriptive information regarding the students taught by ELTs and non-ELTs. Again, consistent with prior work, ELTs tend to be assigned to classrooms with more Black and Hispanic students, more economically disadvantaged students, and students with lower prior test scores and nontest outcomes.[13] For example, classrooms of non-ELTs are 35% economically disadvantaged on average compared with 57% for ELTs and 47% for non-Emergency-license novice teachers post-pandemic. The typical ELT teaches students with lower prior math and ELA achievement on the order of about 0.30 standard deviations.

Appendix Table B3 explores MTEL licensure testing results by licensure status (EL and non-EL) and race/ethnicity. Unsurprisingly, teachers who entered the profession via a route that

[12] We display differences across non-ELT cohorts in Appendix Table B2. Differences across these cohorts are far less pronounced.

[13] We used the Economically Disadvantaged measure constructed by the state. This measure stopped being used by the state following 2021 but is still available in the administrative data for use by researchers to maintain consistency over time.

did not initially require licensure scores tend to score lower on these tests. For example, initial scores by White non-EL holders on MTEL subject tests were over 0.60 standard deviations higher than White Emergency license holders. Even larger gaps emerge for Black and Hispanic teachers.[14] Within the set of ELTs, there are also racial gaps in MTEL performance. For example, White ELTs passed 65% of their initial takes compared with 41% for Black ELTs and 49% for Hispanic ELTs. These results are consistent with prior work that finds ethnoracial gaps in licensure test scores in the state (e.g., Cowan et al., 2023).

5. Empirical Strategy

5.1 Teacher Outcomes

Our first analysis concerns retention in the teaching workforce and evaluation ratings. For both outcomes, we estimate regression models of the form

$$Y_{\mathrm{jt}} = \beta_0 + \beta_1 X_{jt} + \beta_2 ELT_j + \beta_3 Exp_{jt} + \emptyset_{sy} + \varepsilon_{jst} \ , \qquad (1)$$

in which the model predicts outcome Y (retention or evaluating rating) for teacher j in year t as a function of the experience of the teacher Exp_{jt} characteristics of the teacher's school and classroom (X_{jt}), and a school-by-year fixed effect $\emptyset_{sy}$. We estimate two types of attrition: remaining in the same school and remaining in the teaching workforce in the state. We include controls for school fixed effects and classroom characteristics (e.g., student race, economic disadvantage, etc.) to make comparisons across similar teaching contexts. The primary object of interest is β_2, which informs the likelihood of, for example, Emergency license holders remaining in the workforce relative to other teachers with the same experience in similar classroom and school contexts. In these analyses, we also evaluate the extent to which outcomes

[14] Appendix Table B3 displays teachers—rather than all Emergency licenses granted—because the race information used in this paper is obtained from the employment files, not the licensure files. The overall gap between Emergency license and non-Emergency license test scores for all takers is 0.79 standard deviations for first-time subject tests and 0.61 for first-time Communication and Literacy Skills.

differ among candidates from diverse backgrounds. For retention, we pursue these analyses by including interaction terms between a candidate's background (e.g., race/ethnicity) with the ELT indicator in Equation 1.

The inclusion of school-by-year fixed effects in Equation 1 is intended to account for school context by comparing ELTs to other teachers within the same school and year rather than to all teachers in the state. These specifications identify the effects of ELTs relative to teachers who are in similar school settings. This also addresses a possible concern about ELTs possibly being disproportionately hired into schools with a higher degree of learning loss related to the pandemic. We also restrict the sample to post-pandemic-onset years (2021 and later) in our main specifications in order to allow the estimates to be solely obtained from this sample, although results are similar when also including pre-pandemic years.

5.2 Teacher Effectiveness Analyses

We provide a comparison of teacher effectiveness by ELT status by estimating a value-added model of the form

$$y_{ijst} = \beta_0 + \beta_1 X_{ijst} + \beta_2 ELT_j + \beta_3 Exp_{jt} + \emptyset_{sy} + \varepsilon_{ist} \,, \tag{2}$$

in which Y_{ijst} is a student outcome (test scores, nontest factor, or climate views), X_{ijst} is a set of student and classroom characteristics (including prior test scores and nontest outcomes), Exp_{jt} is a vector of teacher experience bins, and $\emptyset_{sy}$ a school-by-year fixed effect. We control for a cubic function of prior test scores in math and ELA, a cubic function of the prior nontest factor, student race, gender, economic disadvantage, limited English proficiency, special education, and class-level averages of each. We also control for bins of teacher experience and grade-by-subject-by-year fixed effects, and we include subject interactions for the prior test and nontest measures along with teacher experience. The regressor of interest, ELT_j, represents whether teacher j entered with an Emergency license, and β_2 thus represents the change in a student or teacher outcome associated with being in a classroom taught by a teacher entering on an Emergency license relative to the combined group of teachers entering through any other route. These methods are standard in the research literature on teacher effectiveness (Chetty et al., 2014a; Jackson, 2018) and routes into the profession (Backes et al., 2018; Backes & Hansen, 2018).

The use of science test scores presents a pair of complications. First, science testing is limited to a subset of grades: Grades 5 and 8 for end-of-grade testing and two end-of-course tests in Grades 9 and/or 10 (biology and physics). We thus cannot control for prior-year science tests as is standard with math and ELA. However, prior math and ELA scores are highly predictive of science scores (Appendix Table B4), suggesting that it is possible that prior math and ELA capture student-teacher sorting in science like they do for math and ELA (Chetty et al., 2014a). We also note that the explained variance (R-squared) of science tests is similar to that of math and ELA, both overall and within school-year. Second, the end-of-course tests are not tied to a specific grade, and some students in Grade 9, for example, take biology, and others take physics. We thus included grade-by-science test (end-of-grade, biology, or physics) interactions.

6. Results

6.1 Main Results

We first display results for retention in Table 3 using within-school-by-year comparisons. In Panel A, we find that ELTs are modestly more likely to be retained in the overall teaching workforce than comparable non–Emergency licensed teachers and to remain in the same school. These differences are quite small, with 0.02 percentage points more likely to stay being about one-fifth of the difference between first- and second-year teachers for both outcomes. In Panel B, we see that these differences are driven by more recent license cohorts. When examining differences by race in Panel C, we find that while Black teachers are less likely to remain in the same school and in the teaching workforce overall, Black teachers on an Emergency license are about as likely to stay as White teachers. Finally, Hispanic teachers with an Emergency license are more likely to remain in the same school than White teachers.

Table 4 displays the main results for estimates of teacher impacts on student outcomes. Column (1) displays a sparse model that only controls for student, classroom, and school characteristics and prior test and nontest outcomes along with teacher experience. In this model, test scores and the nontest factor are negative and statistically significant. Adding school fixed effects (Column [2]) yields estimated changes in the nontest factor that are very small and statistically indistinguishable from 0. This pattern of attenuation with school fixed effects and the differences in classroom composition (Table 2) suggest that some of the displayed negative impacts in Column (1), especially for nontest outcomes, are driven by school context rather than impacts of ELTs themselves.

In Column (3) we include school-by-year fixed effects instead of school fixed effects. Results are very similar, with the exception of results for science being more negative. Columns (4) and (5) repeat these two school effects specifications in our preferred sample of post-pandemic-onset years, 2021 through 2023. Estimates for math and science test scores are consistently negative, and they are statistically significant in the specifications that include school-by-subject-by-year fixed effects (Columns [3] and [5]). One specification that yields an attenuated coefficient in math is the inclusion of school fixed effects in 2021 and after (Column [4]). This is likely due to the dynamics of pandemic-related learning loss in the disadvantaged schools where ELTs are concentrated. In particular, in a school fixed effects model with only three school years, 2020-21 – the first year testing resumed after the onset of the pandemic – receives one-third of the weight. Results for ELTs thus look more rosy when more of the comparison set is drawn from 2020-21. When we include school-by-year fixed effects to only compare within a given year (Columns [3] and [5]), or estimate school fixed effects based on additional data (Column [2]), the remaining estimates are quite consistent.

Table 5 displays results for educator evaluation ratings using the school-year fixed-effect model as in Column (5) of Table 4. Panel A displays overall results for ELTs. Beginning with Column (1), controlling for teacher experience and the characteristics of assigned students, ELTs are 5 percentage points less likely to receive a Proficient or Exemplary rating. This is about the difference in likelihood between a teacher in their first or second year of teaching. In addition, as 93% of recent non–Emergency license recipients receive a Proficient or Exemplary rating (Table 1), these are very large changes in terms of the likelihood of receiving a lower rating. Panel B examines evaluation ratings by license cohort using the cohort definitions defined in Section 4 above. In addition, ELTs have lower evaluation ratings across each of the four standards. Turning to Panel B, all ELT cohorts are less likely to receive a Proficient or Exemplary rating, with the point estimates ranging from 0.04 percent points lower (Cohort 1) to 0.08 percentage points lower (Cohort 3). We find that the differences are largest for Standard I (Curriculum, Planning, and Assessment) and Standard II (Teaching All Students), although the last cohort also has seen very low results for Standard IV (Professional Culture). In Panel C, we examine differences by ELT race. There are few significant differences.

6.2 Heterogeneity by Grade, Subject, and Cohort

We plot grade-by-grade results in Figure 3. The estimated negative effects for math are driven by Grades 4–6. In science, the only tested grade in elementary school is also negative and significant, in addition to negative point estimates in Grade 9 (a mix of biology and introductory physics) and Grade 10 (almost entirely biology).

Figure 4 disaggregates test score results by Emergency license cohort, with cohorts defined by the calendar year in which a teacher first earned their emergency license (January−December). In math, the estimated negative impacts are driven by the cohorts after the first cohort, who earned their first Emergency license between June 2020 and December 2020. When disaggregating effects by both cohort and school level (Appendix Table A2), we find negative effects in math in elementary grades to the order of about 0.040 standard deviations for both Cohorts 2 and 3. This represents a large effect about the difference between a teacher in their first year and fifth or sixth year of teaching (Appendix Table B4). To provide further perspective, 0.040 represents about one-quarter of the total standard deviation of teacher effects for elementary math in Chetty et al. (2014a). Chetty et al. (2014b) also find that a 1 standard deviation increase in teacher quality in one year leads to an increase of $39,000 in lifetime earnings per student. Taking the ELT coefficients in math at face value, this would thus translate to nearly $10,000 in lifetime earnings per student.[15]

Motivated by the observational differences across cohorts in prior preparation and teaching experience, we further explore differences across cohorts in Table 6. Perhaps surprisingly, while there are observable differences across cohorts of ELTs (e.g., Table 1), prior preparation experience, prior employment in Massachusetts schools, and prior MTEL scores do not appear to explain the differences across cohorts in value added in math.[16] One possibility is

[15] Put another way, this represents about three weeks of learning in math (Lipsey et al., 2012).
[16] In Appendix Table A9, we explore other additional factors that could explain cohort differences over time, such as overall changes in the effectiveness of new teachers by including cohort calendar year fixed effects for both ELTs and non-ELTs and possible changes in the composition of the non-ELT comparison group by including controls for initial non-ELT license type. The basic patterns are quite similar to those observed in Table 7.

that teachers who earned their Emergency license in the first few months of the pandemic were disproportionately more likely to be interested in teaching, motivated to obtain the license, and sufficiently organized to do it quickly in a way that cannot be captured by our regressors. An alternative explanation is that subsequent cohorts of ELTs might be more strongly sorted on unmeasured content knowledge if there were changes in the signaling value of the Emergency license. Alternatively, preparation programs may have become less likely to counsel out lower performing students once tests were not required to earn a teaching license.

Finally, we explore differences across cohorts on test score impacts by conversion to a non−Emergency license and Emergency license type in Table 7. To this point, we have used an indicator for whether a teacher earned an Emergency license as their first teaching license as a regressor. This would count, for example, years in which someone was teaching prior to receiving their Emergency license (a possibility if the individual, for example, earned a teaching waiver), introducing potential bias. We thus perform a robustness check in which we only count teachers who had already earned their Emergency license as of a given school year. Because the sample contains data from 2020-21 onward, results for the first cohort are identical because we have already dropped all school years in which the cohort's teachers could have been teaching prior to earning their Emergency license (i.e., the 2020 ELT cohort had already earned their Emergency license as of the 2020−21 school year by definition). For Cohort 3, the results are attenuated, though imprecise, because we had fewer years of available data for this more recent cohort. In the following row, we look at teachers teaching on an Emergency license who earned their Emergency license and no other teaching license. The results are nearly identical. This may not come as a surprise, given that for most cohorts, not enough time passed for meaningful numbers of teachers to have converted to another license during the sample period. Finally, aside from the first cohort, ELTs who eventually convert are less effective than the overall average. This pattern will be worth revisiting as more years of data become available and more ELTs convert. The bottom part of Table 7 examines test score impacts by Emergency license field for the most held fields aligned with available test scores (i.e., not physical education, English as a Second Language, History, etc.). Consistent with the results by subject and school level, the negative effects are concentrated for ELTs in the Elementary and Biology fields.

6.3 Heterogeneity by Student and Teacher Race

ELTs in Massachusetts were more likely to be Black or Hispanic and to teach Black or Hispanic students than teachers entering on Initial or Preliminary licenses. We thus explore the extent to which ELT impacts on student outcomes differ by teacher and by student race. In Table 8, we interact ELT by student race and/or teacher race to obtain race-specific effects for students of color and teachers of color using the school-by-year specification from Column (5) of Table 4. For student race, we find that Black and Hispanic students are impacted similarly to the overall effect: −0.013 standard deviations in Column (5) of Table 4 compared with −0.012 for Black and −0.011 for Hispanic students. For nontest outcomes, we find negative impacts of ELTs on Black students and no statistically significant effects of ELT status on Hispanic students.[17] We also do not find any statistically significant differences in school climate, but results are somewhat imprecisely estimated due to the restricted set of grades covered (Grades 4, 5, 8, and 10). Turning to teacher race, we find that the negative effects on test scores from ELTs are smaller for Hispanic teachers, although results are imprecise due to limited sample sizes of non-White ELTs. Finally, students assigned to Hispanic ELTs tend to have more favorable views of school climate.

[17] The negative impacts on Black students for the nontest factor are driven by students in elementary school; please see Appendix Table A1.

7. Discussion and Conclusion

Early results from Massachusetts and New Jersey have prompted calls to reexamine the role of licensure. For example, in response to early results from Massachusetts and New Jersey, Yglesias (2024) argued that "emergency measures adopted in many states to recruit additional teachers during the pandemic provide further evidence for something many analysts have long believed: Many of the current teacher training and licensing requirements have no real benefits, and getting rid of a lot of them would save time and money for various stakeholders and expand the potential supply of teachers, without reducing quality". The evidence on Emergency licenses presented in this paper suggests early applicants may not be representative of potential candidates when these pathways are more permanent.

In this paper, we provide descriptive evidence on the effectiveness of a larger group of ELTs. One of the challenges in inferring the effects of licensure policies using data on employed teachers is the fact that teachers self-select into licensure pathways. As our signaling model suggests, this implies that comparisons of teachers entering on Emergency and non-Emergency licenses are not necessarily representative of teachers who would have entered only with the change in licensure policy. Moreover, changes in teaching effectiveness across cohorts could plausibly be driven by changes in the selection decisions rather than in the composition of teachers. If later cohorts of ELTs have fewer incumbent teachers, then these cohorts may be better representative of the group of newly employed teachers, but this is difficult to infer from descriptive evidence alone.

Policy effects can vary over time. The positive early findings from Bacher-Hicks et al. (2023) and Backes and Goldhaber (2023) spoke to the impact of early Emergency license recipients. However, compared to subsequent Emergency license cohorts, this initial group was more likely to (at least partially) have met training and licensing requirements. Subsequent ELTs represent a different pool of teachers, one that appears to be less effective in math and science, and rated less effective on the state's educator evaluation system. There are two important caveats to these findings. First, the pool of ELTs includes both newly eligible teachers and teachers who would have obtained employment in the absence of the Emergency license policy. The effectiveness of newly eligible teachers may differ from the average ELT. Second, some of the negative effects of ELTs may be offset if the policy ameliorated teacher shortages that would have been worse if not for the existence of the availability of Emergency licenses. Pandemic-era policy experiments hold important lessons, and both of these remain important issues for future research. However, our results suggest caution about projecting early results into the future. If leaders make pandemic-era policies permanent in a post-pandemic world, they should be prepared for the possibility of unintended consequences and uncertain results.

References

Angrist, J. D., & Guryan, J. (2004). Teacher testing, teacher education, and teacher characteristics. *American Economic Review*, 94(2), 241–246.

Angrist, J. D., & Guryan, J. (2008). Does teacher testing raise teacher quality? Evidence from state certification requirements. *Economics of Education Review*, 27(5), 483–503.

Aldeman, C. (2024a). Emergency-hired teachers do just as well as those who go through normal training. *The74*. Available from https://www.the74million.org/article/emergency-hired-teachers-do-just-as-well-as-those-who-go-through-normal-training/

Aldeman, C. (2024b). Where would you set the bar on the teaching profession? *Aldeman on Education.* Available from https://www.chadaldeman.com/p/where-would-you-set-the-bar-on-the

Bacher-Hicks, A., Chi, O. L., Tichnor-Wagner, A., & Baloch, S. (2023). Teacher Licensure and Workforce Quality: Insights from Covid-Era Emergency Licenses in Massachusetts. *WEPC Working Paper 2024-1*. https://wheelockpolicycenter.org/wp-content/uploads/2024/04/MA-Emergency-License-Teacher-Quality-WP.pdf

Backes, B., Cowan, J., Goldhaber, D., & Theobald, R. (2022). *Teachers and school climate: Effects on student outcomes and academic disparities*. CALDER Working Paper 274–1022.

Backes, B., Cowan, J., Goldhaber, D., & Theobald, R. (2023). How to measure a teacher: The influence of test and nontest value-added on long-run student outcomes. Working Paper No. 270-0423-2. *National Center for Analysis of Longitudinal Data in Education Research (CALDER)*.

Backes, B., & Goldhaber, D. (2023). The relationship between pandemic-era teacher licensure waivers and teacher demographics, retention, and effectiveness in New Jersey. *CALDER Working Paper No. 286-0623*

Backes, B., Goldhaber, D., Cade, W., Sullivan, K., & Dodson, M. (2018). Can Uteach? Assessing the relative effectiveness of STEM teachers. *Economics of Education Review*, *64*, 184–198. https://doi.org/10.1016/j.econedurev.2018.05.002

Backes, B., & Hansen, M. (2018). The impact of Teach For America on non-test academic outcomes. *Education Finance and Policy*, *13*(2), 168–193. http://dx.doi.org/10.1162/EDFP_a_00231

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014a). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, *104*(9), 2593–2632. https://www.aeaweb.org/articles?id=10.1257/aer.104.9.2593

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014b). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American economic review*, *104*(9), 2633-2679.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of education review*, 26(6), 673-682.

Cowan, J., Goldhaber, D., Jin, Z., & Theobald, R. (2023). Assessing Licensure Test Performance and Predictive Validity for Different Teacher Subgroups. *American Educational Research Journal*, *60*(6), 1095-1138.

Cowan, J., Goldhaber, D., & Theobald, R. (2020). *Performance evaluations as a measure of teacher effectiveness when standards differ: Accounting for variation across classrooms, schools, and districts* (No. 140618). National Center for Analysis of Longitudinal Data in Education Research.

DeArmond, M., Goldhaber, D., & Payne, S. (2023). COVID's Under-the-Radar Experiment with Teacher Licensure. CALDER Research Brief. *National Center for Analysis of Longitudinal Data in Education Research (CALDER)*.

Gershenson, S. (2016). Linking teacher quality, student attendance, and student achievement. *Education Finance and Policy*, *11*(2), 125–149. https://doi.org/10.1162/EDFP_a_00180

Gershenson, S., Hart, C. M. D., Hyman, J., Lindsay, C. A., & Papageorge, N. W. (2022). The Long-Run Impacts of Same-Race Teachers. *American Economic Journal: Economic Policy, 14*(4), 300–342. https://doi.org/10.1257/pol.20190573

Goldhaber, D. (2011). Licensure: Exploring the value of this gateway to the teacher workforce. In *Handbook of the Economics of Education* (Vol. 3, pp. 315–339). Elsevier.

Goldhaber, D., & Hansen, M. (2010). Race, gender, and teacher testing: How informative a tool is teacher licensure testing? *American Educational Research Journal*, 47(1), 218-251.

Goodman, G., Arbona, C., & Dominguez de Rameriz, R. (2008). High-stakes, minimum-competency exams: How competent are they for evaluating teacher competence?. *Journal of Teacher Education*, 59(1), 24-39.

Jackson, C. K. (2018). What do test scores miss? The importance of teacher effects on non–test score outcomes. *Journal of Political Economy*, *126*(5), 2072–2107. https://doi.org/10.1086/699018

Lambert, D. (2020, January 23). *California moves closer to eliminating, replacing reading instruction test that has blocked thousands from teaching credential*. EdSource. https://edsource.org/2020/california-moves-closer-to-eliminating-replacing-reading-instruction-test-that-has-blocked-thousands-from-teaching-credential/622830

Partelow, L., Spong, A., Brown, C., & Johnson, S. (2017). America needs more teachers of color and a more selective teaching profession. Washington, DC: *Center for American Progress*, 51.

Podgursky, M. (2005). Teacher licensing in US public schools: The case for simplicity and flexibility. Peabody Journal of Education, 80(3), 15-43.

Nettles, M. T., Scatton, L. H., Steinberg, J. H., & Tyler, L. L. (2011). Performance and passing rate differences of African American and white prospective teachers on PraxisTM examinations: a joint project of the National Education Association (NEA) and Educational Testing Service (ETS). ETS Research Report Series, 2011(1), i-82.

Skinner, K., Betz, K. D., & Wright, A. (2020, April 23). Suspension of teacher license test amid COVID-19 crisis likely to "open up some doors" for potential educators. *Mississippi Today*. https://mississippitoday.org/2020/04/23/suspension-of-teacher-license-test-amid-covid-19-crisis-likely-to-open-up-some-doors-for-potential-educators/

Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics*, *87*(3), 355-374.

Yglesias, M. (2024). The pandemic's lesson on teacher licensure. *Slow Boring*. Available from https://www.slowboring.com/p/the-pandemics-lesson-on-teacher-licensure

Figures and Tables

Figure 1. Teacher Licensure Decision

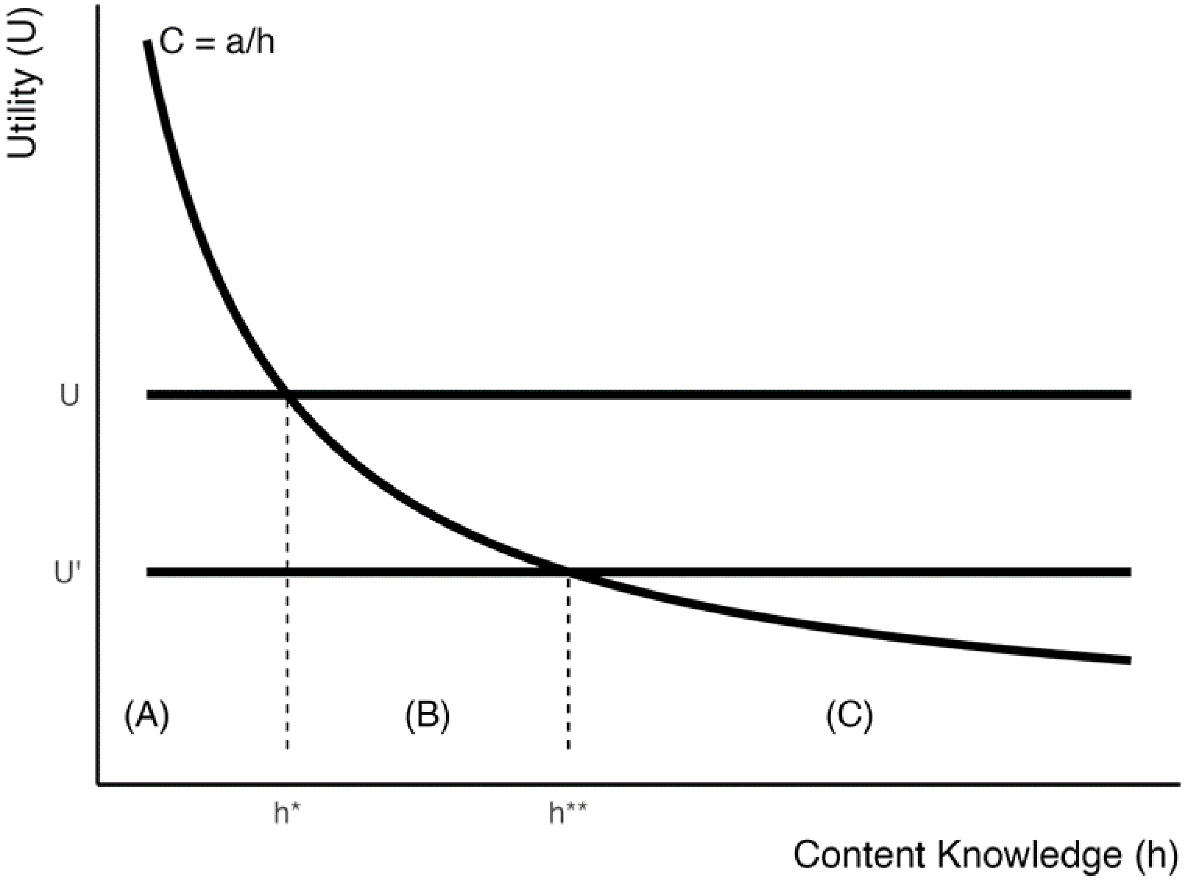


Figure 2. Emergency License Dates

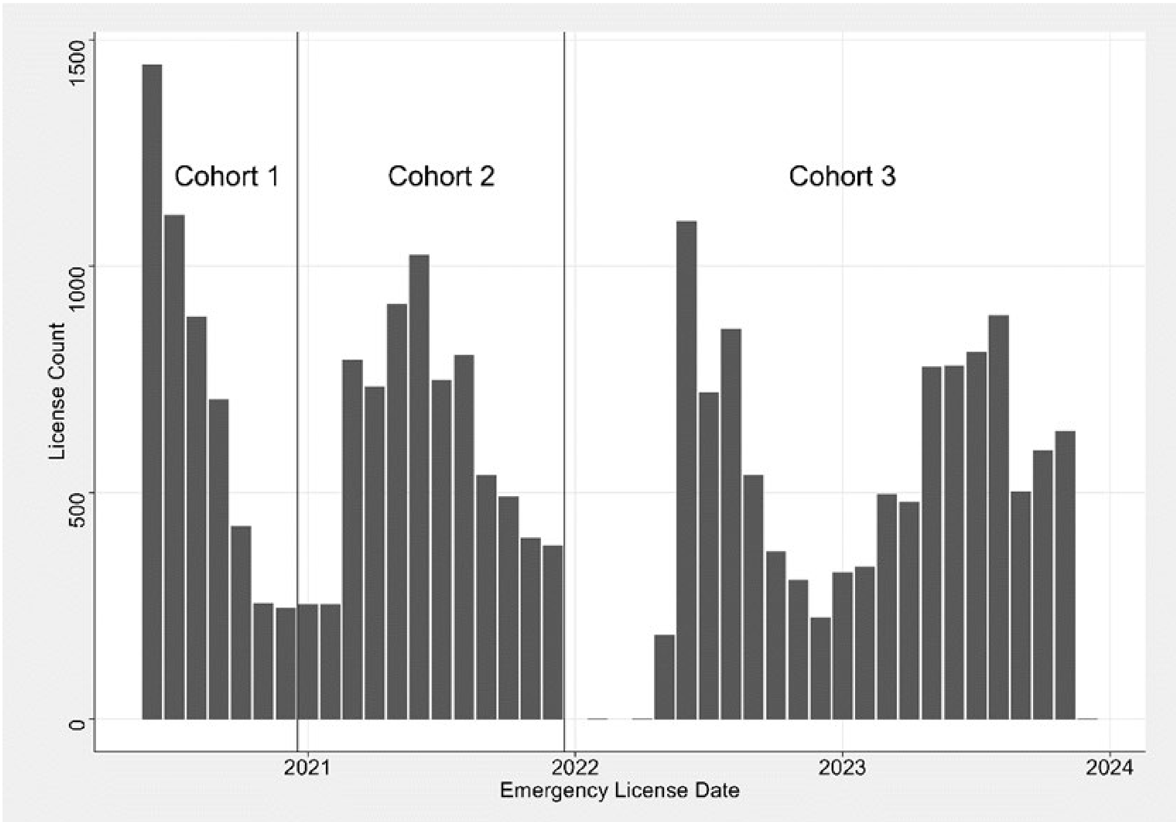
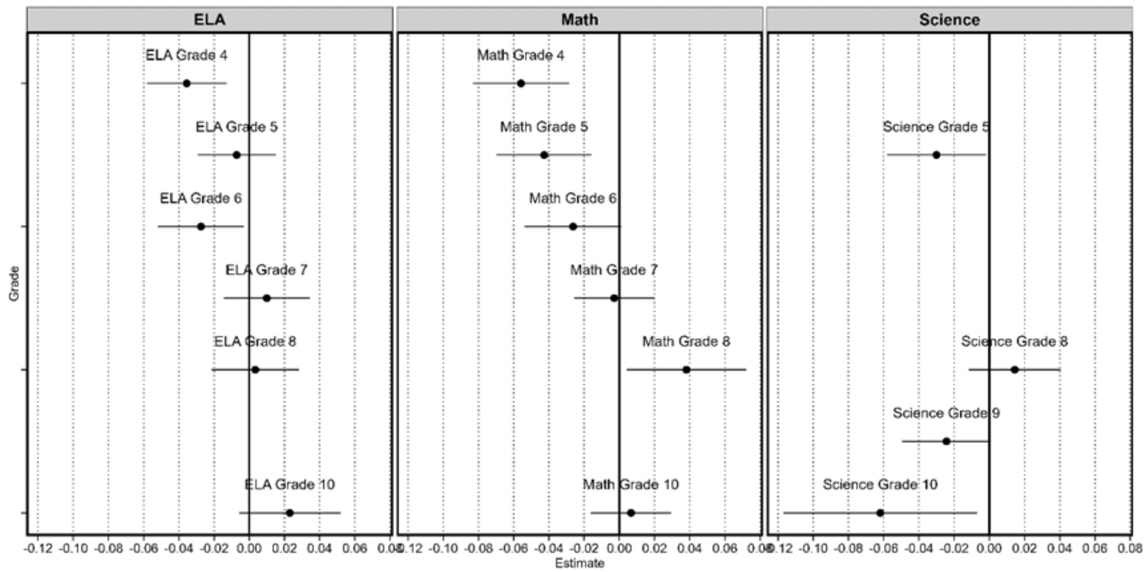
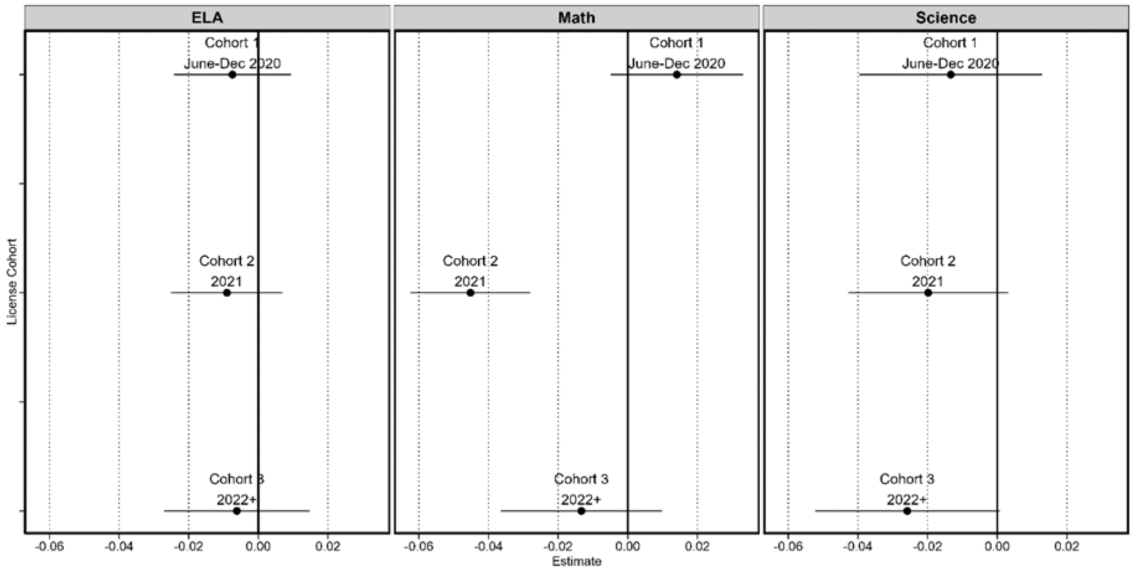


Figure 3. Test Scores by Grade



Notes. Each panel shows the results from a regression of test scores in a given subject interacting grade and Emergency license using the school-by-year fixed-effects model with student and classroom controls shown in Column (5) of Table 4.

Figure 4. Test Scores by Emergency License Cohort



Notes: Each panel shows the results from a regression of test scores in that subject on cohort-by-emergency-license interaction using the school-by-year fixed-effect model with student and classroom controls shown in Column (5) of Table 4. Cohort 1 represents teachers who received their first Emergency license during the 2020 calendar year, Cohort 2 during the 2021 calendar year, etc.

Table 1. Emergency-Licensed Teachers: Comparison Across Cohorts

	(1)	(2)	(3)	(4)
			Emergency	
	Recent Non-ELT	Cohort 1 (2020)	Cohort 2 (2021)	Cohort 3 (2022+)
Ever Core Subj Assignment	0.80	0.74	0.73	0.71
Ever Core Subj SWD Assignment	0.09	0.17	0.12	0.07
Ever Core Subj EL Assignment	0.04	0.06	0.05	0.03
Took MTEL CLST before license	0.93	0.68	0.57	0.43
Took MTEL Subj before license	0.92	0.53	0.42	0.33
Passed MTEL CLST before license	0.93	0.62	0.52	0.38
Passed MTEL Subj before license	0.92	0.27	0.25	0.18
MTEL CLST Std Score	0.28	-0.30	-0.22	-0.33
MTEL Subj Std Score	0.29	-0.49	-0.36	-0.47
Enrolled in prep prior to license	0.57	0.23	0.23	0.14
Teaching position prior to license	0.11	0.24	0.13	0.14
Para prior to license	0.16	0.34	0.26	0.22
Tch White	0.89	0.70	0.73	0.72
Tch Black	0.03	0.14	0.11	0.12
Tch Hispanic	0.04	0.12	0.12	0.12
Tch Female	0.76	0.75	0.74	0.70
Avg evaluation rating	-0.34	-0.56	-0.66	-0.77
Proficient+	0.93	0.87	0.84	0.79
Retained in school	0.69	0.62	0.61	0.67
Retained in district	0.73	0.67	0.67	0.72
Retained in MA	0.82	0.75	0.74	0.80
Observations (unique teachers)	19,290	2,587	3,034	4,187
Total Emergency licenses granted		3,762	4,956	7,241

Notes: Displays Emergency-licensed teachers relative to teachers whose first teaching license was not an Emergency license and who earned their first teaching license in 2017 or later. For Emergency license holders, Columns (2)–(5) show teachers based on the school year during which they earned their first Emergency license. Data include individuals who held a teaching job at some point during the 2020–21 through 2022–23 school years. Bottom row shows total number of Emergency licenses granted, whether or not individual entered teacher labor force.

Table 2. Student Summary Statistics by Teacher License Type

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Novice teachers				
			Non-ELT		Emergency license teachers		
	All	Veteran	Pre	Post	All	Coh. 1	Coh. 2–3
Male	0.52	0.52	0.53	0.53	0.54	0.55	0.54
Hispanic	0.22	0.20	0.27	0.29	0.38	0.37	0.38
Black	0.09	0.09	0.12	0.12	0.14	0.13	0.14
Economic Disadv.	0.35	0.34	0.39	0.47	0.57	0.57	0.57
Limited English Proficiency	0.12	0.11	0.14	0.15	0.20	0.20	0.20
SPED Full Inclusion	0.14	0.14	0.14	0.15	0.14	0.14	0.15
SPED Partial Inclusion	0.04	0.04	0.04	0.04	0.04	0.05	0.04
SPED Substantially Separate	0.04	0.04	0.06	0.06	0.09	0.11	0.08
Secondary STEM class	0.12	0.12	0.12	0.11	0.12	0.12	0.12
Tch-yr obs (linked to student sample)	572,313	480,617	51,627	40,069	16,973	4,689	12,284
Stu Prior Math score	0.01	0.04	-0.13	-0.11	-0.31	-0.33	-0.29
	(0.61)	(0.61)	(0.63)	(0.61)	(0.62)	(0.64)	(0.61)
Stu Prior ELA score	0.01	0.04	-0.13	-0.09	-0.28	-0.31	-0.27
	(0.60)	(0.59)	(0.63)	(0.59)	(0.60)	(0.63)	(0.59)
Stu prior nontest	0.01	0.03	-0.12	-0.07	-0.19	-0.20	-0.18
	(0.65)	(0.64)	(0.76)	(0.61)	(0.63)	(0.64)	(0.63)
Tch-yr obs (prior test sample)	428,404	358,144	45,008	25,252	9,248	3,410	5,838

Notes: Neither test scores nor performance ratings are observed in 2019–20; nontest outcomes not observed in 2020–21. Novice teacher: first, second, or third year of teaching. Veteran teacher: fourth or greater year of teaching. Table 2 shows teachers who enter workforce and are linked to students. Pre-pandemic: school years through 2019–20. Prior student outcomes normalized to be mean 0 in this Table 2 sample.

Table 3. Teacher Retention by ELT Status

	(1) Retained in State	(2) Retained in School
Panel A: Overall		
ELT	0.02***	0.02***
	(0.00)	(0.00)
Panel B: By Emergency License Cohort		
ELT Cohort 1 (2020)	0.01*	0.00
	(0.01)	(0.01)
ELT Cohort 2 (2021)	0.02***	0.03***
	(0.01)	(0.01)
ELT Cohort 3 (2022+)	0.05***	0.05***
	(0.01)	(0.01)
Panel C: By Teacher Race		
ELT	0.02***	0.01*
	(0.00)	(0.01)
Black	-0.05***	-0.04***
	(0.01)	(0.01)
Hispanic	-0.02***	-0.01*
	(0.00)	(0.01)
Black # ELT	0.04***	0.05***
	(0.01)	(0.01)
Hispanic # ELT	0.03**	0.05***
	(0.01)	(0.01)
Observations	224,441	224,441

Notes: Estimates of retention after the 2020–21 through 2022–23 school years with school-by-year fixed effects. Each column represents the results of a regression of a teacher retention outcome on the variables denoted in the first column along with controls for years of experience. Additional controls include cubic polynomial of means at the classroom level of student gender, race, economic disadvantage, English proficiency, and disability status. Standard errors clustered at the teacher level.

Table 4. Estimates of Emergency Licensure on and Student Outcomes

	(1)	(2)	(3)	(4)	(5)
Test Scores	-0.014***	-0.010**	-0.013***	-0.005	-0.012***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Math	-0.023***	-0.017***	-0.015**	-0.007	-0.015**
	(0.008)	(0.006)	(0.006)	(0.007)	(0.007)
ELA	0.000	0.002	-0.005	0.003	-0.008
	(0.007)	(0.006)	(0.005)	(0.006)	(0.006)
Science	-0.017	-0.013	-0.025***	-0.018*	-0.018**
	(0.012)	(0.010)	(0.009)	(0.009)	(0.009)
Nontest	-0.021***	-0.003	-0.006	-0.003	-0.007
Factor	(0.007)	(0.006)	(0.005)	(0.006)	(0.005)
Log Abs	0.023***	0.010**	0.000	0.000	0.001
	(0.006)	(0.004)	(0.003)	(0.004)	(0.004)
Log Susp	0.009***	0.003	0.003*	0.001	0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Retained	0.000	0.000	0.001*	0.001	0.001
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
GPA	-0.011	0.005	-0.001	0.018***	-0.001
(Secondary)	(0.010)	(0.007)	(0.006)	(0.007)	(0.006)
VOCAL	0.011	0.019*	0.011	0.006	0.010
	(0.012)	(0.010)	(0.009)	(0.010)	(0.010)
School ctrl	Yes				
Sch-Subj FE		Yes		Yes	
Sch-Subj-Yr FE			Yes		Yes
2021+ only				Yes	Yes

Notes: Regression of outcome on cubic functions of prior test scores in math and ELA, a cubic function of the prior nontest factor, student race, gender, economic disadvantage, limited English proficiency, special education, and class- level averages of each. Column (1) additionally controls for each of these averages at the school level. Regression also controls for bins of teacher experience, grade-by-subject-by-year fixed effects, and subject interactions for the prior test and nontest measures along with teacher experience. Standard errors clustered at the teacher level.

Table 5. Educator Evaluation Ratings for Emergency License Teachers

			By Standard (STD)			
	Proficient+	Overall	STD I	STD II	STD III	STD IV
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Overall						
ELT	-0.05***	-0.19***	-0.18***	-0.16***	-0.11***	-0.15***
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Panel B: By Emergency License Cohort						
ELT Cohort 1 (2020)	-0.04***	-0.16***	-0.17***	-0.14***	-0.09***	-0.12***
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
ELT Cohort 2 (2021)	-0.05***	-0.21***	-0.18***	-0.17***	-0.13***	-0.17***
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
ELT Cohort 3 (2022+)	-0.08***	-0.25***	-0.21***	-0.19***	-0.12***	-0.19***
	(0.01)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)
Panel C: By Teacher Race						
ELT	-0.04***	-0.17***	-0.15***	-0.14***	-0.09***	-0.12***
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Black	-0.03***	-0.19***	-0.19***	-0.16***	-0.14***	-0.20***
	(0.00)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Hispanic	-0.01**	-0.10***	-0.09***	-0.08***	-0.05**	-0.11***
	(0.00)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Black # ELT	-0.03**	-0.03	-0.04	-0.02	-0.08*	-0.05
	(0.01)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Hispanic # ELT	-0.01	0.01	0.01	0.01	0.05	0.00
	(0.01)	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)
Observations	196,843	196,843	196,843	196,843	196,843	196,843

Notes: Teacher-level regression with school-by-year fixed effects and controls for years of experience along with cubic polynomial of means at the classroom level of student gender, race, economic disadvantage, English proficiency, and disability status. Standard errors clustered at the teacher level. The outcome variable in the Proficient+ column is an indicator of whether the teacher was rated as Proficient or Exemplary, and the remaining columns use scores standardized to be mean 0, standard deviation 1 within each school year.

Standard I: Curriculum, Planning, and Assessment
Standard II: Teaching All Students
Standard III: Family and Community Engagement
Standard IV: Professional Culture
Information: https://www.doe.mass.edu/edeval/model/PartIII_AppxC.pdf

Table 6. Estimates of Emergency License Holder Effectiveness by Prior Measures

	Math (1)	Math (2)	Math (3)	Science (4)	Science (5)	Science (6)
ELT	0.014	0.044***	0.059***	-0.013	-0.017	-0.018
	(0.010)	(0.014)	(0.016)	(0.013)	(0.018)	(0.021)
ELT x Cohort 2	-0.059***	-0.062***	-0.063***	-0.007	-0.006	-0.013
	(0.012)	(0.012)	(0.013)	(0.017)	(0.017)	(0.019)
ELT x Cohort 3	-0.027*	-0.034**	-0.052***	-0.013	-0.009	-0.022
	(0.014)	(0.014)	(0.017)	(0.017)	(0.018)	(0.021)
ELT x Prior Prep		0.010	0.010		0.026	0.028
		(0.013)	(0.014)		(0.017)	(0.019)
ELT x Prior Empl.		0.002	0.006		-0.013	0.006
		(0.011)	(0.013)		(0.015)	(0.019)
ELT x Took MTEL		-0.045***	-0.050***		0.003	0.001
		(0.012)	(0.015)		(0.015)	(0.020)
MTEL Subj Score			0.012***			0.010***
			(0.002)			(0.003)
Observations	1,019,121	1,019,121	895,731	473,396	473,396	413,663

Notes: Regression of test score outcome on cubic functions of prior test scores in math and ELA, a cubic function of the prior nontest factor, student race, gender, economic disadvantage, limited English proficiency, special education, and class-level averages of each. Regressions also controls for bins of teacher experience, grade-by-year fixed effects, and teacher experience. Standard errors clustered at the teacher level. Prior prep indicates whether a teacher enrolled in a teacher preparation program prior to receiving their first teaching license. Prior employment denotes whether a teacher was employed in MA public schools prior to earning their teaching license. MTEL Subj. Score represents standardized MTEL scores on the subject license test.

Table 7. Test Score Impacts by Cohort and Sample

	Cohort 1 (2020)	Cohort 2 (2021)	Cohort 3 (2022+)
Overall	0.002	-0.026***	-0.012
	(0.006)	(0.006)	(0.008)
By Active EL and Conversion			
On Emergency License	0.002	-0.026***	-0.010
	(0.006)	(0.006)	(0.008)
On EL, has not converted	0.000	-0.022***	-0.010
	(0.007)	(0.006)	(0.008)
Ever convert EL	0.015*	-0.027***	-0.016
	(0.008)	(0.007)	(0.015)
By Emergency License Type			
Elementary	-0.011	-0.047***	-0.031**
	(0.011)	(0.009)	(0.013)
Moderate Disabilities	0.004	-0.011	0.040
	(0.012)	(0.013)	(0.024)
English	-0.014	0.010	0.008
	(0.012)	(0.012)	(0.013)
Mathematics	0.025*	-0.029**	0.019
	(0.013)	(0.012)	(0.014)
General Science	0.019	-0.036*	0.006
	(0.026)	(0.019)	(0.024)
Biology	-0.049**	-0.049**	-0.012
	(0.020)	(0.021)	(0.021)
Middle School Math/Sci	0.007	-0.020	0.013
	(0.019)	(0.022)	(0.027)

Notes: Each row represents a separate regression of Emergency license cohort interacted with the variable indicated in the first column plus cubic functions of prior test scores in math and ELA, a cubic function of the prior nontest factor, student race, gender, economic disadvantage, limited English proficiency, special education, and class-level averages of each. Regressions also controls for bins of teacher experience, grade-by-year fixed effects, and teacher experience. Standard errors clustered at the teacher level. "On Emergency License" represents teachers who earned Emergency license as of current school year (i.e., discards observations from ELTs in the years teaching prior to earning license). This is identical to the overall results from Cohorts 1 and 2 because the sample excludes school years prior to 2020–21. "On EL, has not converted" represents teachers who have earned an Emergency license and not yet converted their license to a non–Emergency license as of given school year. Emergency license type regressions interact cohort with an indicator for whether an individual ever earned an Emergency license in a specific field; these categories are not mutually exclusive.

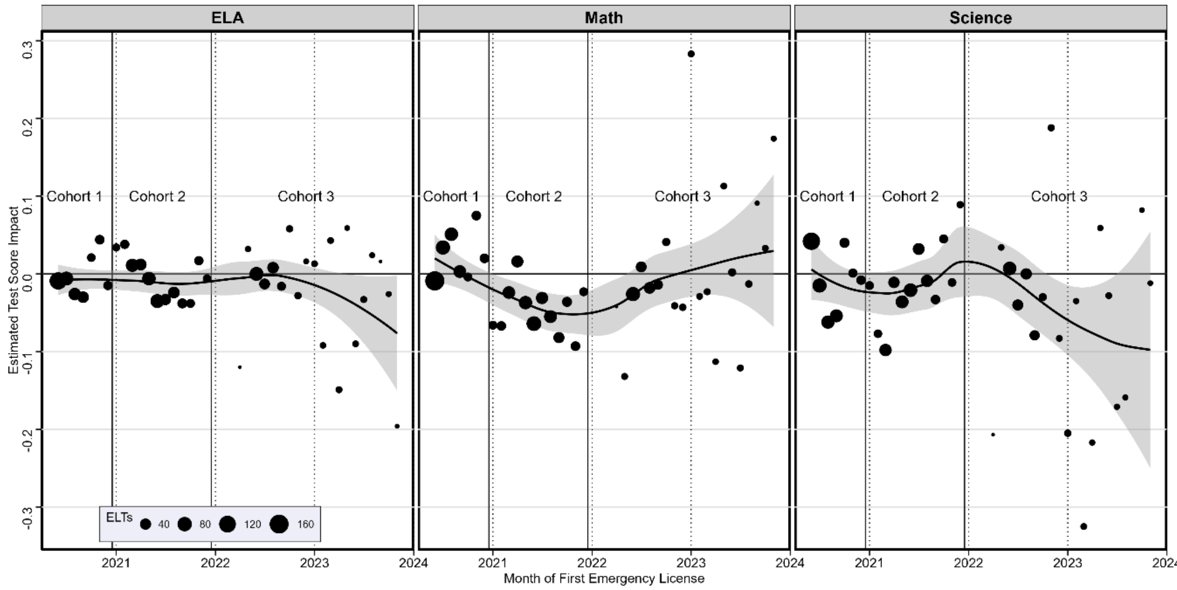
Table 8. Heterogeneous Student Impacts for Teachers and Student of Color

	Test	Test	Nontest	Nontest	VOCAL	VOCAL
ELT x Student Race						
Black	-0.012*		-0.028**		0.014	
	(0.006)		(0.012)		(0.018)	
Hispanic	-0.011**		0.003		-0.004	
	(0.005)		(0.008)		(0.013)	
ELT x Teacher Race						
Black		-0.013		-0.005		0.007
		(0.010)		(0.016)		(0.027)
Hispanic		-0.008		-0.004		0.057**
		(0.010)		(0.015)		(0.028)

Notes: Regression of outcome denoted by column header on cubic functions of prior test scores in math and ELA, a cubic function of the prior nontest factor, student race, gender, economic disadvantage, limited English proficiency, special education, and class-level averages of each. Regression also controls for bins of teacher experience, grade-by-subject-by-year fixed effects, and subject interactions for the prior test and nontest measures along with teacher experience. Each regression includes the school-by-subject-by-year fixed effects used in Column (5) of Table 4. Standard errors clustered at the teacher level. Teacher main effects for race not included. Omitted group are students taught by non-ELTs.

Appendix A. Additional Results

Figure A1. Month by ELT Interactions



Notes: Estimates obtained from interacting Emergency license with month of first Emergency license receipt using the school-by-year fixed effect model in Column (5) of Table 4. Smoothed line weights each point by the number of Emergency license recipients in each month bin.

Table A1. Heterogeneity by Student Race and School Level

	Test	Nontest	Log Absence	Log Days Susp	VOCAL
Panel 1: Elementary School					
ELT × Stu Black	-0.016	-0.059**	-0.001	0.014**	0.016
	(0.011)	(0.030)	(0.017)	(0.006)	(0.029)
ELT × Stu Hispanic	-0.026***	-0.007	-0.019*	0.004	-0.003
	(0.009)	(0.017)	(0.011)	(0.004)	(0.020)
Panel 2: Middle School					
ELT × Stu Black	-0.011	-0.009	0.011	0.015**	0.050**
	(0.009)	(0.016)	(0.010)	(0.007)	(0.023)
ELT × Stu Hispanic	-0.003	0.014	-0.011	0.001	-0.002
	(0.008)	(0.012)	(0.007)	(0.004)	(0.017)
Panel 3: High School					
ELT × Stu Black	0.014	0.003	-0.016	0.021**	0.001
	(0.014)	(0.014)	(0.012)	(0.010)	(0.030)
ELT × Stu Hispanic	0.001	0.007	-0.012	-0.006	-0.025
	(0.010)	(0.009)	(0.011)	(0.005)	(0.021)

Notes: Regression of outcome in a given subsample on cubic functions of prior test scores in math and ELA, a cubic function of the prior nontest factor, student race, gender, economic disadvantage, limited English proficiency, special education, and class-level averages of each. Regression also controls for bins of teacher experience and grade-by-year fixed effects. Each regression includes school-by-year fixed effects. Standard errors clustered at the teacher level.

Table A2. Emergency License Holder Effectiveness by Test Subject, Level, and Cohort

	ELA (1)	Math (2)	Science (3)
Panel 1: Elementary School			
Cohort 1 (2020)	-0.010	-0.004	-0.038*
	(0.014)	(0.016)	(0.022)
Cohort 2 (2021)	-0.016	-0.042***	-0.015
	(0.011)	(0.013)	(0.017)
Cohort 3 (2022+)	-0.009	-0.033*	-0.047**
	(0.018)	(0.019)	(0.019)
Panel 2: Middle School			
Cohort 1 (2020)	-0.009	0.011	-0.002
	(0.012)	(0.014)	(0.021)
Cohort 2 (2021)	0.009	-0.049***	-0.023
	(0.012)	(0.013)	(0.016)
Cohort 3 (2022+)	0.002	0.003	-0.012
	(0.016)	(0.016)	(0.018)
Panel 3: High School			
Cohort 1 (2020)	0.032	0.012	-0.007
	(0.021)	(0.016)	(0.019)
Cohort 2 (2021)	0.014	0.003	-0.020
	(0.017)	(0.014)	(0.020)
Cohort 3 (2022+)	-0.035	0.009	-0.011
	(0.024)	(0.019)	(0.023)
School-Yr FE	Yes	Yes	Yes

Notes: Regression of outcome in a given subsample on cubic functions of prior test scores in math and ELA, a cubic function of the prior nontest factor, student race, gender, economic disadvantage, limited English proficiency, special education, and class-level averages of each. Regression also controls for bins of teacher experience and grade-by-year fixed effects. Each regression includes the school-by-year fixed effects. Standard errors clustered at the teacher level. Cohort defined by calendar year in which first teaching license was received.

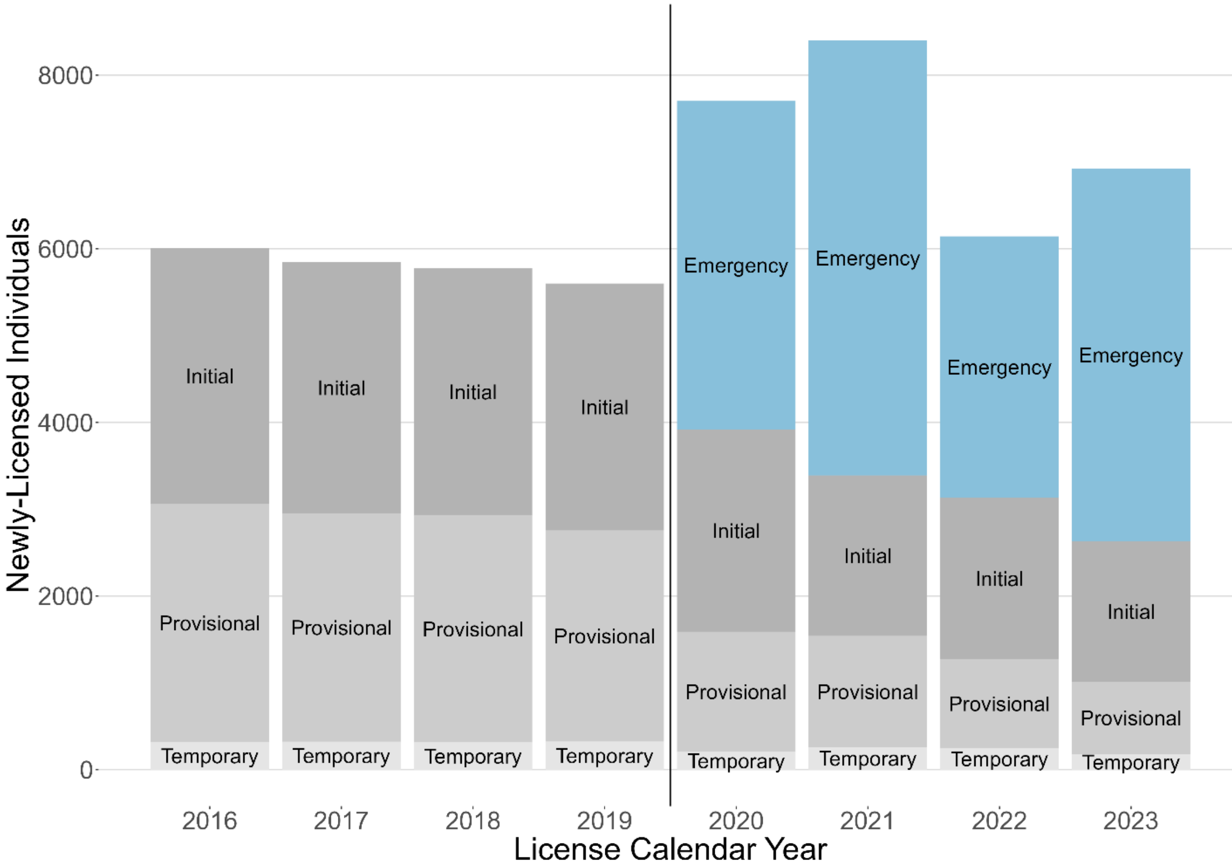
Table A3. Differences Across License Cohorts and License Types

	ELA	ELA	ELA	Math	Math	Math	Science	Science	Science
By EL Cohort									
1 (2020)	-0.007	-0.012	-0.011	0.014	0.013	0.010	-0.013	-0.029*	-0.032*
	(0.009)	(0.011)	(0.011)	(0.010)	(0.013)	(0.013)	(0.013)	(0.016)	(0.017)
2 (2021)	-0.009	-0.005	0.000	-0.045***	-0.035***	-0.039***	-0.020*	-0.006	-0.010
	(0.008)	(0.011)	(0.011)	(0.009)	(0.013)	(0.014)	(0.012)	(0.016)	(0.016)
3 (2022+)	-0.006	0.010	0.006	-0.013	-0.006	-0.012	-0.026*	-0.010	-0.011
	(0.011)	(0.015)	(0.015)	(0.012)	(0.016)	(0.017)	(0.014)	(0.022)	(0.021)
Post-Pandemic License Calendar Year FE									
2020		0.003	0.006		-0.001	-0.002		0.015	0.012
		(0.008)	(0.008)		(0.010)	(0.011)		(0.011)	(0.012)
2021		-0.007	-0.004		-0.013	-0.016		-0.016	-0.018
		(0.009)	(0.010)		(0.012)	(0.012)		(0.014)	(0.014)
2022		-0.014	-0.008		-0.013	-0.020		-0.025	-0.032*
		(0.012)	(0.012)		(0.013)	(0.014)		(0.020)	(0.019)
2023		-0.048**	-0.036		-0.005	-0.014		0.006	0.003
		(0.023)	(0.023)		(0.024)	(0.025)		(0.041)	(0.031)
First License									
Prov.			-0.002			-0.002			0.001
			(0.004)			(0.004)			(0.005)
Temp.			-0.008			0.000			0.024
			(0.010)			(0.010)			(0.015)
N	1,012,819	1,012,819	804,952	1,019,121	1,019,121	821,127	473,396	473,396	374,535

Notes: Regression of test score outcome on cubic functions of prior test scores in math and ELA, a cubic function of the prior nontest factor, student race, gender, economic disadvantage, limited English proficiency, special education, and class-level averages of each. Regressions also controls for bins of teacher experience, grade-by-year fixed effects, and teacher experience. Standard errors clustered at the teacher level. Cohort indicators denote calendar year of initial license receipt

Appendix B. Additional Summary Statistics and Data

Figure B1. Newly-Licensed Individuals by License Type



Notes: Year indicates calendar year in which first teaching license was received.

Table B1. Emergency License Counts

Description	**Unique Individuals**
Received Emergency license	20,710
Emergency license is a teaching license	16,719
No prior teaching license	15,959
Employed after obtaining Emergency license	11,625
Held classroom (teacher or para) job after Emergency license	11,010
Held teaching job classification after Emergency license	9,875
Emergency license and in test score outcome sample (Table 4)	1,988

Table B2. Non-ELT Summary Statistics by License Cohort

	(1) 2017–19	(2) 2020	(3) 2021	(4) 2022	(5) 2023
First license Initial	0.49	0.60	0.57	0.58	0.62
First license Provisional	0.43	0.34	0.32	0.32	0.31
First license Temporary	0.04	0.04	0.06	0.07	0.06
Took MTEL CLST before license	0.94	0.95	0.90	0.90	0.91
Took MTEL Subj before license	0.93	0.94	0.89	0.90	0.89
Passed MTEL CLST before license	0.94	0.95	0.90	0.90	0.91
Passed MTEL Subj before license	0.93	0.94	0.89	0.90	0.88
MTEL CLST Std Score	0.27	0.28	0.29	0.32	0.27
MTEL Subj Std Score	0.27	0.28	0.35	0.39	0.33
Enrolled in prep prior to license	0.55	0.64	0.58	0.58	0.59
Completed prep prior to license	0.43	0.54	0.50	0.51	0.49
Held teaching position prior to license	0.10	0.11	0.11	0.10	0.13
Para prior to license	0.18	0.17	0.14	0.12	0.12
Tch White	0.89	0.89	0.87	0.88	0.88
Tch Black	0.03	0.03	0.03	0.03	0.02
Tch Hispanic	0.04	0.04	0.04	0.04	0.05
Tch Female	0.76	0.78	0.74	0.73	0.76
Std Eval	-0.27	-0.40	-0.50	-0.58	-0.51
Retained in school	0.71	0.66	0.67	0.68	0.56
Retained in district	0.74	0.70	0.70	0.72	0.62
Retained in MA	0.82	0.81	0.81	0.86	0.82
Observations	10,342	2,517	2,124	1,925	1,304

Notes: By calendar year of first license receipt.

Table B3. First-Time MTEL results by Teacher Race and License Type

	(1)	(2)	(3)	(4)	(5)	(6)
	Non−Emergency License			Emergency License		
	White	Black	Hispanic	White	Black	Hispanic
Took MTEL CLST	0.99	0.98	0.99	0.93	0.93	0.90
Took MTEL Subj.	0.98	0.98	0.98	0.83	0.73	0.77
MTEL CLST score	0.30	0.01	0.05	-0.13	-0.77	-0.76
	(0.68)	(0.76)	(0.79)	(0.86)	(1.00)	(1.10)
MTEL Subj. score	0.31	-0.18	0.11	-0.32	-1.15	-0.63
	(0.80)	(0.83)	(0.82)	(1.02)	(1.02)	(1.11)
Pct. First-time tests passed	0.83	0.71	0.76	0.65	0.41	0.49
Ever fail a MTEL test	0.34	0.53	0.43	0.55	0.76	0.69
Pct. Initial failures retaken	0.58	0.53	0.55	0.43	0.28	0.30
Tests taken within	2.93	2.87	2.88	2.55	2.19	2.18
9 mo. Of initial test	(1.12)	(1.13)	(1.04)	(1.08)	(1.01)	(0.97)
Avg. times taking each test	1.14	1.20	1.17	1.20	1.21	1.19
	(0.32)	(0.37)	(0.32)	(0.41)	(0.43)	(0.41)
Unique Teachers	16,526	561	828	5478	904	907

Notes: Sample consists of teachers who earned their first license in 2017 or later and held a teaching position in at least one year from 2020–21 through 2022–23. MTEL scores standardized in full MTEL sample, which includes teachers and non-teachers.

Table B4. Main Specification Full Regression Coefficients

	ELA	Math	Science
Emergency license	-0.008	-0.015**	-0.018**
	(0.006)	(0.007)	(0.009)
2nd Year of Teaching	0.014**	0.022***	0.002
	(0.006)	(0.006)	(0.008)
3rd/4th Year of Teaching	0.019***	0.036***	0.015*
	(0.006)	(0.007)	(0.008)
5th/6th Year of Teaching	0.039***	0.049***	0.032***
	(0.006)	(0.007)	(0.009)
7th–10th Year of Teaching	0.044***	0.070***	0.030***
	(0.006)	(0.007)	(0.008)
Over 10 Years of Teaching	0.055***	0.083***	0.045***
	(0.005)	(0.006)	(0.007)
Prior Math	0.257***	0.690***	0.465***
	(0.002)	(0.002)	(0.003)
Prior Math ^ 2	0.023***	0.045***	0.042***
	(0.001)	(0.001)	(0.001)
Prior Math ^ 3	-0.008***	-0.037***	-0.019***
	(0.000)	(0.000)	(0.001)
Prior ELA	0.576***	0.159***	0.400***
	(0.002)	(0.001)	(0.002)
Prior ELA ^ 2	-0.011***	-0.006***	-0.015***
	(0.001)	(0.001)	(0.001)
Prior ELA ^ 3	-0.031***	-0.006***	-0.024***
	(0.000)	(0.000)	(0.001)
Prior Nontest Factor	0.013***	0.041***	0.040***
	(0.001)	(0.001)	(0.002)
Prior Nontest Factor ^ 2	0.000	0.005***	0.004***
	(0.001)	(0.001)	(0.001)
Prior Nontest Factor ^ 3	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Male	-0.124***	0.082***	0.127***
	(0.001)	(0.001)	(0.002)
Hispanic	-0.017***	-0.037***	-0.062***
	(0.002)	(0.002)	(0.003)
Black	-0.017***	-0.075***	-0.102***
	(0.003)	(0.002)	(0.004)
Asian	0.023***	0.137***	-0.017***
	(0.003)	(0.003)	(0.004)
HPI	0.012	0.020	-0.029
	(0.019)	(0.017)	(0.026)
American Indian	-0.023**	-0.017	-0.036**
	(0.012)	(0.011)	(0.017)

Multiple races	0.012***	0.002	0.004
	(0.003)	(0.003)	(0.004)
Economic Disadvantage	-0.055***	-0.061***	-0.056***
	(0.001)	(0.001)	(0.002)
SPED Full Inclusion	-0.182***	-0.150***	-0.072***
	(0.002)	(0.002)	(0.003)
SPED Partial Inclusion	-0.298***	-0.280***	-0.185***
	(0.006)	(0.006)	(0.007)
SPED Substantially Separate	-0.321***	-0.334***	-0.195***
	(0.013)	(0.013)	(0.017)
Limited English Proficiency	-0.177***	-0.081***	-0.130***
	(0.003)	(0.003)	(0.005)
Grade 9 * Physics			0.002
			(3.626)
Grade 9 * Biology			0.159
			(3.628)
Grade 10 * Biology			0.145***
			(0.043)
Observations	1,012,819	1,019,121	473,396
R2	0.67	0.74	0.69
Within R2	0.59	0.66	0.60

Notes: Full regression coefficients for Column (5) of Table 4. Regression also controls for grade-by-year fixed effects, school-by-year fixed effects, and classroom averages of each of the covariates shown. Standard errors clustered at the teacher level.